

Appendix B:

Probability of reproducing 1A7 in a second hybridoma from a different animal

Background:

<p>Hapten survey</p> <p><i>(Appendix A of this document, outlining current understanding of the generation of antibody diversity, and illustrated by articles provided herewith)</i></p>	<ul style="list-style-type: none">• Multiple VJ gene choices and splices are allowed to form suitable prototype light chains to bind any particular antigen.• Multiple VDJ gene choices and splices are allowed to form suitable prototype heavy chains to bind any particular antigen.• At least ~83% of the amino acids can vary from prototype sequences.• At least ~4 variations are allowed on average in the other positions. (The 12 sequences obtained by Leahy et al. (<i>Appendix A</i>) have an average of ~3.5 variations in the 83% of the positions that vary. 12 is a small sample size for such a variable systems, and these numbers are therefore underestimates. On average, any codon can mutate to codons for 6 distinct amino acids with a single base change.)• A plurality of mutation events may be involved in creating antigen binding activity. However, it is rare that a particular amino acid mutation at a particular position is absolutely required in order to create antigen binding activity. Alternatives are permitted at all other locations.• Mutation events are not antigen driven, they are antigen selected. Mutations may occur throughout the variable region. Mutations not deleterious to antigen binding are tolerated. Accordingly, many mutations in a heavily mutated sequence (particularly those outside the antigen binding site, which is generally comprised within the CDRs) are irrelevant to antigen binding, and hence to clonal selection.
<p>1A7 analysis</p> <p><i>(Exhibit B. of the Sunil Chatterjee declaration)</i></p>	<ul style="list-style-type: none">• 1A7 light chain variable region has at least 2 point differences from the prototype VJ sequence.• 1A7 heavy chain variable region has at least 14 point differences from the prototype VDJ sequence.• 9 of these 16 point differences occur <i>outside</i> the complementarity determining regions (CDRs). The 7 other differences are broadly distributed amongst the CDRs.

The observations above and the calculations that follow are all based on data for antibody variable region *amino acid* sequences. (Possible “non-productive” mutations in the ambiguous positions of the corresponding DNA sequence are neither included nor relied upon.)

Probability calculation using conservative assumptions

- Assume only a small fraction of available V_H , D_H , J_H , V_L and J_L genes are suitable for generating an antibody specific for the immunizing antigen.
- Assume only a small fraction of splice options are suitable.
- Assume the expected distribution of the number of point mutations obtained under immunization conditions used is narrow, and number of mutations in 1A7 is typical.
- Assume as many as 17% of the amino acids may not be varied from the prototype heavy and light chain sequences. In addition, assume as many as 3 of the 14 point mutations from the heavy chain prototype sequence are absolutely and invariably required in order to bind antigen.
- Assume the number of differences allowed at each of the other point mutations is only 4 per position on average.
- Assume only kappa light chain (not lambda chain) is suitable.

<i>Occurrence of a cell producing a 1A7 V_H region (117 residues)</i>			
• Probability of obtaining exactly 14 mutations (Depends on mean and distribution of expected number of mutations under immunization conditions used)	<i>1 in</i>		5
• Possible number of acceptable point differences at each mutation point	<i>1 in</i>	$4^{11} \times 1^3 =$	4.2×10^3
• Possible locations of the 11 non-mandatory differences	<i>1 in</i>	$\frac{(117 \times 83\%)!}{(117 \times 83\% - 11)! 11!} =$ $\frac{94!}{83! 11!} =$	6.9×10^{13}
• V_H gene selection	<i>1 in</i>		5
• D_H gene selection	<i>1 in</i>		4
• J_H gene selection	<i>1 in</i>		2
• Splice options	<i>1 in</i>		10
Total Compound Probability	<i>1 in</i>		5.8×10^{23}

Occurrence of a cell producing a 1A7 V_L region (112 residues)

• Probability of obtaining exactly 2 mutations under immunization conditions used	1 in		5
• Possible number of acceptable point differences	1 in	$4^2 =$	16
• Possible locations of the 2 differences	1 in	$\frac{(112 \times 83\%)!}{(112 \times 83\% - 2)! 2!} =$ $\frac{96!}{94! 2!} =$	4.6×10^3
• V _H gene selection	1 in		5
• J _H gene selection	1 in		2
• Splice options	1 in		3
Total Compound Probability	1 in		1.1×10^7
<i>Occurrence of a cell producing an entire 1A7 variable region</i>		1 in	6.3×10^{30}

Probability calculation using non-conservative assumptions

- Assume a moderate fraction of available V_H , D_H , J_H , V_L and J_L genes are suitable for generating an antibody specific for the immunizing antigen.
- Assume a moderate fraction of splice options are suitable.
- Assume the expected distribution of the number of point mutations obtained under immunization conditions used is moderate, and number of mutations in 1A7 is typical.
- Assume 3 amino acids may not be varied in heavy and light chain sequences (two C and one W) that are invariant between immunoglobulin domains. Assume 1 amino acid in the heavy chain is absolutely required for antigen binding.
- Assume the number of differences allowed elsewhere is 20 per position.

<i>Occurrence of a cell producing a 1A7 V_H region (117 residues)</i>			
• Probability of obtaining exactly 14 mutations under immunization conditions used (Bernoulli distribution; depends on standard deviation)	<i>1 in</i>		50
• Possible number of acceptable point differences	<i>1 in</i>	$20^{14} =$	1.6×10^{18}
• Possible locations of the differences	<i>1 in</i>	$\frac{(117 - 4)!}{(117 - 4 - 14)!14!} =$	2.7×10^{17}
• Possible non-repetitive VDJ selection and splice combinations	<i>1 in</i>		10000
Total Compound Probability	<i>1 in</i>		2.2×10^{41}
<i>Occurrence of a cell producing a 1A7 V_L region (112 residues)</i>			
• Probability of obtaining exactly 2 mutations under immunization conditions used	<i>1 in</i>		50
• Possible number of acceptable point differences	<i>1 in</i>	$20^2 =$	400
• Possible locations of the 2 differences	<i>1 in</i>	$\frac{(112 - 3)!}{(112 - 3 - 2)!2!} =$	5.9×10^3
• Possible non-repetitive VJ selection and splice combinations	<i>1 in</i>		1000
• Choice of kappa or lambda chain	<i>1 in</i>		2
Total Compound Probability	<i>1 in</i>		2.4×10^{11}
<i>Occurrence of a cell producing an entire 1A7 variable region</i>			<i>1 in</i>
			5.3×10^{52}

Number of mice required to regenerate 1A7 antibody

It is difficult to estimate the total number of somatically mutated antibodies that arise from a single primary B cell. Patten et al. have suggested that the upper bound for the number of variants occurring during each immunization is 10^6 . Multiple immunizations could theoretically increase this number. Some mutations, however, would render the subsequent lineage either unable to form a properly folded immunoglobulin, or would ablate antigen-binding activity.

A much clearer limit is set by the number of suitable antibody-producing cells present in the mouse at the time of harvesting. An ordinary mouse spleen has between 10^8 and 10^9 cells, the majority of which will not be specific for antigen. According to the Declaration by Malaya Bhattacharia-Chatterjee, only 1 well out of 1024 was positive when fused cells from 4 immunized mice were plated. This indicates that generally no more than a few (certainly not more than 10^3) antibody-producing cells can be identified, fused and expanded from each immunized mouse according to the protocol used.

Using the conservative estimate *supra* for the number of possible antibody molecules capable of binding the immunizing antigen, the frequency of a second mouse harboring a spleen cell making antibody identical to 1A7 is no more than:

$$1 \text{ in } \frac{6 \times 10^{30}}{10^3} = 6 \times 10^{27} \text{ mice}$$